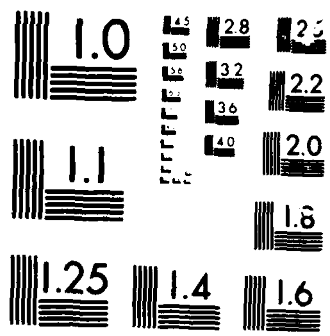
AD-A167 534 THE CENTRAL LIMIT THEOREM AND POINCARE-TYPE 1/1
INEQUALITIES(U) WISCONSIN UNIV-MADISON MATHEMATICS
RESEARCH CENTER L H CHEN APR 86 MRC-TSR-2929
UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL



MICROCOP CHART

AD-A167 534

(2)

MRC Technical Summary Report #2929

THE CENTRAL LIMIT THEOREM AND POINCARÉ-TYPE INEQUALITIES

Louis H. Y. Chen

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

April 1986

(Received March 21, 1986)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

THE CENTRAL LIMIT THEOREM AND POINCARÉ-TYPE INEQUALITIES

Louis H. Y. Chen

Technical Summary Report #2929

April 1986

ABSTRACT

We use Poincaré-type inequalities to prove the sufficiency and necessity of the Lindeberg condition in the central limit theorem.

AMS (MOS) Subject Classifications: Primary 60F05, 60E15; Secondary 26D10.

Key Words: Central limit theorem, Poincaré-type inequalities, characterization of the normal distribution

Work Unit Number 4 - Statistics and Probability

Research partially sponsored by the United States Army under Contract No. DAAG29-80-C-0041. Part of this work was done when the author was visiting the Institute of Mathematics and Its Applications at the University of Minnesota.

SIGNIFICANCE AND EXPLANATION

The central limit theorem is a fundamental theorem in probability and statistics. It states that the probability distribution of the sum of a large number of small and mutually independent random numerical observations approaches a normal distribution as the number of observations increases. The Lindeberg condition is a condition for which the central limit theorem holds. It has been proved to be both necessary and sufficient.

A Poincaré-type inequality is an inequality which relates the integral of the square of a function to the integral of the square of its derivative. In this report we give a new proof of the central limit theorem by using Poincaré-type inequalities to prove both the necessity and sufficiency of the Lindeberg condition.

Accesion For		
NTIS CRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

THE CENTRAL LIMIT THEOREM AND POINCARÉ-TYPE INEQUALITIES

Louis H. Y. Chen

Borovkov and Utev (1984) defined the functional

$$U(X) = \sup_g \frac{\mathrm{Var}[g(X)]}{\sigma^2 E[g'(X)]^2}$$

for any random variable X with finite variance σ^2, where the supremum is taken over the class of absolutely continuous functions g such that $0 < \mathrm{Var}[g(X)] < \infty$. They proved that

(1) $U(X) \geq 1$ and if $U(X) = 1$ then X has a normal distribution.

Using this result they further proved that

(2) if $X_1, X_2, \ldots$ is a sequence of random variables such that $U(X_n) \to 1$ then the moment generating function of $(X_n - EX_n)/[\mathrm{Var}(X_n)]^{1/2}$ exists and converges to that of the standard normal random variable in a neighborhood of zero.

It is natural to ask if (1) can also be applied to prove the central limit theorem under the Lindeberg condition. This question was in fact raised by S. Kotani in private communication with the author and motivated the present work.

The existence of the moment generating function of X_n in (2) is due to the finiteness of $U(X_n)$ (see Borovkov and Utev, Theorem 2). Since the central limit theorem does not require such a strong condition, arguments different from those of Borovkov and Utev would have to be used. It turns out that a Poincaré-type inequality for sums of independent random variables proved along the same line as in Chen (1985, Section 2) is the key to this problem. Using this inequality we can prove not only the sufficiency of the

---

Research partially sponsored by the United States Army under Contract No. DAAG29-80-C-0041. Part of this work was done when the author was visiting the Institute for Mathematics and Its Applications at the University of Minnesota.

Lindeberg condition but also its necessity.

In order to facilitate application, we restate (1) in a different form. Let $C_B^1(\mathbb{R})$ be the class of functions g such that g and g' are bounded and continuous.

PROPOSITION 1. Let X be a random variable with finite variance $\sigma^2 > 0$. If $\mathrm{Var}[g(X)] \le \sigma^2 E[g'(X)]^2$ for $g \in C_B^1(\mathbb{R})$, then X has a normal distribution.

To see that (1) and Proposition 1 are equivalent, we define

$$U_0(X) = \sup_{\substack{g \in C_0^\infty(\mathbb{R}) \\ \mathrm{Var}[g(X)]>0}} \frac{\mathrm{Var}[g(X)]}{\sigma^2 E[g'(X)]^2}$$

and

$$U_B(X) = \sup_{\substack{g \in C_B^1(\mathbb{R}) \\ \mathrm{Var}[g(x)]>0}} \frac{\mathrm{Var}[g(X)]}{\sigma^2 E[g'(X)]^2}$$

where $C_0^\infty(\mathbb{R})$ is the class of C^∞ functions on $\mathbb{R}$ with compact support. Then we observe that by definition, $U(X) \ge 1$ and that by Theorem 2(i) of Borovkov and Utev, $U_0(X) = U_B(X) = U(X)$.

Since we are interested in both an application of Poincaré-type inequalities and a new proof of the central limit theorem, it is fitting to examine the arguments which lead to (1) and hence Proposition 1. There are three different proofs of (1). All begin with a variational argument. After that the first uses the method of moments (see Borovkov and Utev, Theorem 3). The second uses the characteristic function (see Chen and Lou (1985, Theorem 2.1 and Corollary 2.1). The third uses differential equations (see Chen and Lou, Lemmas 4.1 and 4.2).

THE CENTRAL LIMIT THEOREM AND POINCARÉ-TYPE INEQUALITIES

Louis H. Y. Chen

Borovkov and Utev (1984) defined the functional

$$U(X) = \sup_g \frac{\mathrm{Var}[g(X)]}{\sigma^2 E[g'(X)]^2}$$

for any random variable X with finite variance σ^2, where the supremum is taken over the class of absolutely continuous functions g such that $0 < \mathrm{Var}[g(X)] < \infty$. They proved that

(1) $U(X) \geq 1$ and if $U(X) = 1$ then X has a normal distribution.

Using this result they further proved that

(2) if $X_1, X_2, \ldots$ is a sequence of random variables such that $U(X_n) \to 1$ then the moment generating function of $(X_n - EX_n)/[\mathrm{Var}(X_n)]^{1/2}$ exists and converges to that of the standard normal random variable in a neighborhood of zero.

It is natural to ask if (1) can also be applied to prove the central limit theorem under the Lindeberg condition. This question was in fact raised by S. Kotani in private communication with the author and motivated the present work.

The existence of the moment generating function of X_n in (2) is due to the finiteness of $U(X_n)$ (see Borovkov and Utev, Theorem 2). Since the central limit theorem does not require such a strong condition, arguments different from those of Borovkov and Utev would have to be used. It turns out that a Poincaré-type inequality for sums of independent random variables proved along the same line as in Chen (1985, Section 2) is the key to this problem. Using this inequality we can prove not only the sufficiency of the

Research partially sponsored by the United States Army under Contract No. DAAG29-80-C-0041. Part of this work was done when the author was visiting the Institute for Mathematics and Its Applications at the University of Minnesota.

Lindeberg condition but also its necessity.

In order to facilitate application, we restate (1) in a different form. Let $C_B^1(\mathbb{R})$ be the class of functions g such that g and g' are bounded and continuous.

PROPOSITION 1. Let X be a random variable with finite variance $\sigma^2 > 0$. If $\mathrm{Var}[g(X)] < \sigma^2 E[g'(X)]^2$ for $g \in C_B^1(\mathbb{R})$, then X has a normal distribution.

To see that (1) and Proposition 1 are equivalent, we define

$$U_0(X) = \sup_{\substack{g \in C_0^\infty(\mathbb{R}) \\ \mathrm{Var}[g(X)]>0}} \frac{\mathrm{Var}[g(X)]}{\sigma^2 E[g'(X)]^2}$$

and

$$U_B(X) = \sup_{\substack{g \in C_B^1(\mathbb{R}) \\ \mathrm{Var}[g(x)]>0}} \frac{\mathrm{Var}[g(X)]}{\sigma^2 E[g'(X)]^2}$$

where $C_0^\infty(\mathbb{R})$ is the class of C^∞ functions on $\mathbb{R}$ with compact support. Then we observe that by definition, $U(X) > 1$ and that by Theorem 2(i) of Borovkov and Utev, $U_0(X) = U_B(X) = U(X)$.

Since we are interested in both an application of Poincaré-type inequalities and a new proof of the central limit theorem, it is fitting to examine the arguments which lead to (1) and hence Proposition 1. There are three different proofs of (1). All begin with a variational argument. After that the first uses the method of moments (see Borovkov and Utev, Theorem 3). The second uses the characteristic function (see Chen and Lou (1985, Theorem 2.1 and Corollary 2.1). The third uses differential equations (see Chen and Lou, Lemmas 4.1 and 4.2).

We now prove the sufficiency of the Lindeberg condition. Let $X_{n1},\ldots,X_{nr_n}$, $n \geq 1$, be a triangular array of row-wise independent random variables with zero means and finite variances $\sigma_{n1}^2,\ldots,\sigma_{nr_n}^2$ such that $\sum_{i=1}^{r_n} \sigma_{ni}^2 = 1$. Let $W_n = \sum_{i=1}^{r_n} X_{ni}$ and $W_n^{(i)} = W_n - X_{ni}$. By Theorem 2.1 of Chen (1985),

$$\mathrm{Var}[g(W_n)] \leq \sum_{i=1}^{r_n} E\Big[\mathrm{Var}^{W_n^{(i)}} g(W_n)\Big]$$

for $g \in C^1(\mathbb{R})$ such that $|g(x)| \leq C(1 + |x|)$ for some constant C. Now the right hand side of the inequality equals

$$\sum_{i=1}^{r_n} FE^{W_n^{(i)}}\big[g(W_n) - g(W_n^{(i)}) - E^{W_n^{(i)}}(g(W_n) - g(W_n^{(i)}))\big]^2$$

$$\leq \sum_{i=1}^{r_n} E[g(W_n) - g(W_n^{(i)})]^2$$

$$= \sum_{i=1}^{r_n} E\Big[\int_0^{X_{ni}} g'(W_n^{(i)} + t)dt\Big]^2$$

$$\leq \sum_{i=1}^{r_n} E\Big\{X_{ni}\int_0^{X_{ni}} [g'(W_n^{(i)} + t)]^2 dt\Big\}$$

$$= \sum_{i=1}^{r_n} E\Big\{X_{ni}\int_{-\infty}^{\infty} [I(X_{ni} > t > 0) - I(X_{ni} < t < 0)][g'(W_n^{(i)} + t)]^2 dt\Big\}$$

$$= \sum_{i=1}^{r_n} \int_{-\infty}^{\infty} E[g'(W_n^{(i)} + t)]^2 K_n^{(i)}(t)dt$$

where $K_n^{(i)}(t) = EX_i[I(X_{ni} > t > 0) - I(X_{ni} < t < 0)] \geq 0$. Define the probability measure ν_n on $\mathcal{B}(\mathbb{R}^2)$ by

$$\int_{\mathbb{R}^2} f d\nu_n = \sum_{i=1}^{r_n} \int_{-\infty}^{\infty} Ef(W_n^{(i)}, t)K_n^{(i)}(t)dt$$

for bounded and continuous functions f on $\mathbb{R}^2$. Also define $\psi : \mathbb{R}^2 \to \mathbb{R}$ by $\psi(x,y) = x+y$. Then, combining the above inequalities, we have the following Poincaré-type inequality for W_n : For $g \in C^1(\mathbb{R})$ such that $|g(x)| \leq C(1 + |x|)$ for some constant C,

$$\text{(3)} \qquad \operatorname{Var}[g(W_n)] \leq \int_{\mathbb{R}} (g')^2 d\nu_n \circ \psi^{-1} \quad .$$

Let $\nu_n^{(1)}(A) = \nu_n(A \times \mathbb{R})$ and $\nu_n^{(2)}(A) = \nu_n(\mathbb{R} \times A)$ for $A \in \mathcal{B}(\mathbb{R})$. Since

$$\begin{aligned} \nu_n^{(2)}(|y| > \varepsilon) &= \sum_{i=1}^{r_n} \int_{|t|>\varepsilon} K_n^{(i)}(t)dt \\ &= \sum_{i=1}^{r_n} E|X_{ni}|(|X_{ni}| - \varepsilon)^+ \\ &\leq \sum_{i=1}^{r_n} EX_{ni}^2 I(|X_{ni}| > \varepsilon) \quad , \end{aligned}$$

the Lindeberg condition implies that $\nu_n^{(2)} \Longrightarrow \varepsilon_0$, the Dirac measure at 0, as $n \to \infty$. Now $EW_n^{(i)2} \leq EW_n^2 = 1$ for each n and i implies that $\{\nu_n^{(1)}\}$ is tight and hence relatively compact. Let $\{\nu_{n'}^{(1)}\}$ be a weakly convergent subsequence and let $\nu_{n'}^{(1)} \Longrightarrow L(Z)$. Then $\nu_{n'} \Longrightarrow L\binom{Z}{0}$ and hence $\nu_{n'} \circ \psi^{-1} \Longrightarrow L(Z)$. It follows from (3) that for $g \in C_B^1(\mathbb{R})$,

$$\text{(4)} \qquad \operatorname{Var}[g(Z)] \leq E[g'(Z)]^2 \quad .$$

If we show that Z has a standard normal distribution, then $L(W_n) \Longrightarrow N(0,1)$ is proved. By Proposition 1 it remains to prove that $\operatorname{Var}(Z) = 1$. First we observe that by virtue of $EW_n^2 = 1$, Z is square integrable. Let $\varphi_a \in C_B^1(\mathbb{R})$ be increasing such that

$$\varphi_a(x) = \begin{cases} x & \text{if } |x| \leq a \\ a+1 & \text{if } x > a+2 \\ -a-1 & \text{if } x < -a-2 \quad . \end{cases}$$

By (3),

$$\operatorname{Var}[W_{n'} - \varphi_a(W_{n'})] \leq \int_{\mathbb{R}} (1-\varphi_a')^2 d\nu_{n'} \circ \psi^{-1} \quad .$$

Since $\nu_{n'} \circ \psi^{-1}$ converges weakly, $\{\nu_{n'} \circ \psi^{-1}\}$ is tight. Therefore for $\varepsilon > 0$,

$$\operatorname{Var}[W_{n'} - \varphi_a(W_{n'})] \leq \int_{\mathbb{R}} (1-\varphi_a')^2 d\nu_{n'} \circ \psi^{-1} \leq \varepsilon$$

for sufficiently large a. Now $(\mathrm{Var})^{1/2}$ is a seminorm and so

$$|1 - [\mathrm{Var}(\varphi_a(W_{n'}))]^{1/2}| = |[\mathrm{Var}(W_{n'})]^{1/2} - [\mathrm{Var}(\varphi_a(W_{n'}))]^{1/2}|$$

$$< [\mathrm{Var}(W_{n'} - \varphi_a(W_{n'}))]^{1/2}$$

$$< \varepsilon^{1/2} \quad .$$

By letting $n' \to \infty$ and then $a \to \infty$, we obtain

$$|1 - [\mathrm{Var}(Z)]^{1/2}| < \varepsilon^{1/2} \quad .$$

This implies that $\mathrm{Var}(Z) = 1$ and hence $L(W_n) \Longrightarrow N(0,1)$.

We now prove the necessity of the Lindeberg condition. First we need two *simple propositions*.

PROPOSITION 2. The Lindeberg condition holds if and only if $\nu_n^{(2)} \Longrightarrow \varepsilon_0$.

PROOF. The "only if" part has been proved above. The "if" part follows from the following inequalities:

$$\sum_{i=1}^{r_n} E|X_{ni}|(|X_{ni}| - \varepsilon)^+ > \sum_{i=1}^{r_n} E|X_{ni}|(|X_{ni}| - \varepsilon)I(|X_{ni}| > 2\varepsilon)$$

$$> \frac{1}{2}\sum_{i=1}^{r_n} EX_{ni}^2 I(|X_{ni}| > 2\varepsilon) \quad .$$

For the next proposition let $C_U^2(\mathbb{R})$ be the class of functions g on $\mathbb{R}$ such that g, g' and g'' are uniformly continuous, $|g(x)| < C(1 + |x|)$ for some constant C and g' and g'' are bounded.

PROPOSITION 3. Let Z and T be independent random variables such that Z has the normal distribution with mean 0 and variance $\sigma^2 > 0$. If $\mathrm{Var}[g(Z)] < \sigma^2 E[g'(Z+T)]^2$ for $g \in C_U^2(\mathbb{R})$, then $T = 0$ w.p. 1.

PROOF. By the variational argument used in Borovkov and Utev (1984) or Chen and Lou (1985), we have

$$EZh(Z) = \sigma^2 Eh'(Z+T)$$

for $h \in C_U^2(\mathbb{R})$. This equation also holds for $h(x) = x^3$ by approximating this function by functions of $C_U^2(\mathbb{R})$. So

$$3\sigma^4 = EZ^4 = 3\sigma^2 E(Z+T)^2 \quad .$$

The finiteness of $E(Z+T)^2$ and EZ^2 implies that of ET^2. By expanding $E(Z+T)^2$ we obtain $\sigma^2 = \sigma^2 + ET^2$ which implies that $ET^2 = 0$ and so $T = 0$ w.p. 1.

For the proof of the necessity of the Lindeberg condition we need the usual assumption that for every $\varepsilon > 0$, $\max_{1 \le i \le r_n} P(|X_{ni}| > \varepsilon) \to 0$ as $n \to \infty$. Let $g \in C_U^2(\mathbb{R})$. For every $\varepsilon > 0$, let $\delta > 0$ be such that $|(g'(x))^2 - (g'(y))^2| < \varepsilon$ for $|x-y| < \delta$. Then we have

$$\left| \sum_{i=1}^{r_n} \int_{-\infty}^{\infty} E\{(g'(W_n^{(i)} + t))^2 - (g'(W_n + t))^2\} K_n^{(i)}(t)dt \right|$$

$$\le C\{ \sum_{i=1}^{r_n} P(|X_{ni}| > \delta) \int_{-\infty}^{\infty} K_n^{(i)}(t)dt + \varepsilon \sum_{i=1}^{r_n} \int_{-\infty}^{\infty} K_n^{(i)}(t)dt\} \tag{5}$$

$$\le C[\max_{1 \le i \le r_n} P(|X_{ni}| > \delta) + \varepsilon]$$

for some constant C. Define the probability measure $\tilde{\nu}_n$ on $\mathcal{B}(\mathbb{R})$ by

$$\int_{\mathbb{R}} f d\tilde{\nu}_n = \sum_{i=1}^{r_n} \int_{-\infty}^{\infty} f(t) K_n^{(i)}(t)dt$$

for bounded and continuous functions f on $\mathbb{R}$. In view of (5), the inequality (3) can be written as

$$\mathrm{Var}[g(W_n)] \le \int_{-\infty}^{\infty} E[g'(W_n+t)]^2 \, \tilde{\nu}_n(dt) + o(1) \quad . \tag{6}$$

Suppose $\mathcal{L}(W_n) \Longrightarrow N(0,1)$. Let $\{\tilde{\nu}_{n'}\}$ be a subsequence of $\{\tilde{\nu}_n\}$ which converges vaguely to a subprobability measure $\tilde{\nu}$. Then for $g \in C_0^2(\mathbb{R})$, the class of C^2 functions on $\mathbb{R}$ with compact support, it is not difficult to show that

$$\int_{-\infty}^{\infty} E[g'(W_n + t)]^2 \, \tilde{\nu}_{n'}(dt) \to \int_{-\infty}^{\infty} E[g'(Z+T)]^2 \, \tilde{\nu}(dt)$$

where Z is a standard normal random variable. So by (6) we have

$$\text{(7)} \qquad \operatorname{Var}[g(Z)] \le \int_{-\infty}^{\infty} E[g'(Z+T)]^2 \, \tilde{\nu}(dt) \quad .$$

By approximating $C_U^2(\mathbb{R})$ by functions of $C_0^2(\mathbb{R})$, (7) holds for $g \in C_U^2(\mathbb{R})$. By letting $g(x) = x$, we get

$$1 \le \int_{-\infty}^{\infty} \tilde{\nu}(dt)$$

and so $\tilde{\nu}$ is a probability measure. By Proposition 3, $\tilde{\nu}$ must be ε_0. Hence $\tilde{\nu}_n \Longrightarrow \varepsilon_0$. But $\tilde{\nu}_n = \nu_n^{(2)}$. By Proposition 2, the Lindeberg condition holds. This completes the proof.

REFERENCES

BOROVKOV, A. A. and UTEV, S. A. (1984). On an inequality and a related characterization of the normal distribution. Theory Prob. Appl. 28, 219-228.

CHEN, L. H. Y. (1985). Poincaré-type inequalities via stochastic integrals. Z. Wahrscheinlichkeitstheorie verw. Gebiete 69, 251-277.

CHEN, L. H. Y. and LOU, J. H. (1985). Characterization of probability distributions by Poincaré-type inequalities. Preprint.

LHYC/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2929	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE CENTRAL LIMIT THEOREM AND POINCARÉ-TYPE INEQUALITIES		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Louis H. Y. Chen		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE April 1986
		13. NUMBER OF PAGES 8
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Central limit theorem, Poincaré-type inequalities, characterization of the normal distribution.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We use Poincaré-type inequalities to prove the sufficiency and necessity of the Lindeberg condition in the central limit theorem.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

6-86

DTIC